Length Frequency and Proportional Size Distribution (PSD) of Largemouth Bass 2014-2016

Alex Benecke

October 20, 2017

```
PSD <- read.csv("Data/Clean-Data/2012-2016_nearshore-survey-largemouth-bass_CLEAN.csv") %>%
  arrange(Year,FID,Length)
PSD$fyr <- as.factor(PSD$fyr)

str(PSD)
```

```
## 'data.frame':    335 obs. of  16 variables:
##  $ Year  : int  2012 2012 2012 2012 2012 2012 2012 2012 2012 2012 ...
##  $ Site  : int  18 18 18 18 18 18 18 18 18 18 ...
##  $ FID   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Weight: num  155 145 170 700 850 750 850 800 950 850 ...
##  $ Length: int  220 220 230 347 364 368 368 371 374 377 ...
##  $ AC    : int  2 3 3 3 3 3 3 3 3 3 ...
##  $ AGE   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ SexCon: int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Sex   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Delts : logi  NA NA NA NA NA NA ...
##  $ logW  : num  2.19 2.16 2.23 2.85 2.93 ...
##  $ logL  : num  2.34 2.34 2.36 2.54 2.56 ...
##  $ fyr   : Factor w/ 4 levels "2012","2014",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ Ws    : num  138 138 159 612 715 ...
##  $ Wr    : num  113 105 107 114 119 ...
##  $ gcat  : Factor w/ 3 levels "preferred","quality",..: 3 3 3 2 2 2 2 2 2 2 ...
```

```
headtail(PSD)
```

```
##     Year  Site FID Weight Length AC AGE SexCon Sex Delts     logW     logL
## 1   2012    18  NA    155    220  2  NA     NA  NA    NA 2.190332 2.342423
## 2   2012    18  NA    145    220  3  NA     NA  NA    NA 2.161368 2.342423
## 3   2012    18  NA    170    230  3  NA     NA  NA    NA 2.230449 2.361728
## 333 2016    15 130    305    266  3   2      8   2    NA 2.484300 2.424882
## 334 2016    15 131    282    261  3   2      3   1    NA 2.450249 2.416641
## 335 2016 15972 132    971    395  3   7      3   1    NA 2.987219 2.596597
##      fyr       Ws       Wr      gcat
## 1   2012 137.6415 112.6114     stock
## 2   2012 137.6415 105.3461     stock
## 3   2012 159.1971 106.7859     stock
## 333 2016 256.2345 119.0316     stock
## 334 2016 240.8044 117.1075     stock
## 335 2016 934.6786 103.8860 preferred
```

```
unique(PSD$Year) ### See that there is no 2013
```

```
## [1] 2012 2014 2015 2016
```

```
PSD %<>% mutate(lcat20=lencat(Length,w=20)); headtail(PSD)
```

```
##     Year  Site FID Weight Length AC AGE SexCon Sex Delts     logW     logL
## 1   2012    18  NA    155    220  2  NA     NA  NA    NA 2.190332 2.342423
## 2   2012    18  NA    145    220  3  NA     NA  NA    NA 2.161368 2.342423
## 3   2012    18  NA    170    230  3  NA     NA  NA    NA 2.230449 2.361728
## 333 2016    15 130    305    266  3   2      8   2    NA 2.484300 2.424882
## 334 2016    15 131    282    261  3   2      3   1    NA 2.450249 2.416641
## 335 2016 15972 132    971    395  3   7      3   1    NA 2.987219 2.596597
##      fyr       Ws       Wr      gcat lcat20
## 1   2012 137.6415 112.6114     stock    220
## 2   2012 137.6415 105.3461     stock    220
## 3   2012 159.1971 106.7859     stock    220
## 333 2016 256.2345 119.0316     stock    260
## 334 2016 240.8044 117.1075     stock    260
## 335 2016 934.6786 103.8860 preferred    380
```

```
### just looking at data
PSD[c(300:335),c(1,3,5,16,17)]
```

```
##     Year FID Length      gcat lcat20
## 300 2016  95    290     stock    280
## 301 2016  96    231     stock    220
## 302 2016  97    298     stock    280
## 303 2016  98    275     stock    260
## 304 2016  99    289     stock    280
## 305 2016 100    279     stock    260
## 306 2016 101    310   quality    300
## 307 2016 102    289     stock    280
## 308 2016 103    336   quality    320
## 309 2016 104    220     stock    220
## 310 2016 105    373   quality    360
## 311 2016 106    254     stock    240
## 312 2016 107    245     stock    240
## 313 2016 108    289     stock    280
## 314 2016 109    358   quality    340
## 315 2016 111    265     stock    260
## 316 2016 112    269     stock    260
## 317 2016 113    336   quality    320
## 318 2016 114    307   quality    300
## 319 2016 115    275     stock    260
## 320 2016 116    358   quality    340
## 321 2016 117    279     stock    260
## 322 2016 118    324   quality    320
## 323 2016 119    236     stock    220
## 324 2016 120    259     stock    240
## 325 2016 122    236     stock    220
## 326 2016 123    288     stock    280
## 327 2016 124    202     stock    200
## 328 2016 125    296     stock    280
## 329 2016 126    269     stock    260
## 330 2016 127    283     stock    280
## 331 2016 128    376   quality    360
## 332 2016 129    362   quality    360
## 333 2016 130    266     stock    260
## 334 2016 131    261     stock    260
## 335 2016 132    395 preferred    380
```

```
psd.12 <- filter(PSD,Year==2012)
psd.13 <- NA
psd.14 <- filter(PSD,Year==2014)
psd.15 <- filter(PSD,Year==2015)
psd.16 <- filter(PSD,Year==2016)
```

My plan is to make a data file with the PSD of Prefered and Quality length largemouth bass for each year. This will be used for comparison graph against Relative weight and should provide some insight into whether the population is expiriencing slow growth (perhaps due to competition for resources) or excessive mortality (perhaps due to overfishing). Both of these things can affect the proportion of large individuals and thus angler satisfaction and may be correctable with regulatory changes.

Possible conclusions about the status of the population (Guy and Brown pg 412).

1) **Poor Habitat** low recruitment, slow growth, and moderate to high mortality due to poor habitat.
2) **Overharvest** of largemouth bass greater than quality length (or maybe prefered).
3) **Stunting** high density of small, slow-growing largemouth bass due to excessive recruitment resulting in stunted growth due to exsesive intraspecific competition at young ages.

Note I am not using data from 2012 because of differences in the data colection procedures. Only large fish were measures weighed and aged all others were just counted.

Note data from 2013 has been removed because I do not have a single data file with lengths and weights together. Although I do have these variables seperatly without any individual identifyer.

Length Frequency

```
### 2014
(psd.14.freq20 <- xtabs(~lcat20,data=psd.14))
```

```
## lcat20
## 200 220 240 260 280 300 320 340 360 380 400 420 460 480
##   1   3  16  28  17  13  10  22  12   6   7   2   2   1
```

```
### calculate the percentage of the fish in each length nin
prop.table(psd.14.freq20)*100
```

```
## lcat20
##        200        220        240        260        280        300
##  0.7142857  2.1428571 11.4285714 20.0000000 12.1428571  9.2857143
##        320        340        360        380        400        420
##  7.1428571 15.7142857  8.5714286  4.2857143  5.0000000  1.4285714
##        460        480
##  1.4285714  0.7142857
```

```
### 2015
(psd.15.freq20 <- xtabs(~lcat20,data=psd.15))
```

```
## lcat20
## 200 220 240 260 280 300 320 340 360 380 400 420 440 460
##   3   2   2   3   4   8  13   8   9   7   4   2   1   1
```

```
### calculate the percentage of the fish in each length nin
prop.table(psd.15.freq20)*100
```

```
## lcat20
```

```
##       200       220       240       260       280       300       320
##  4.477612  2.985075  2.985075  4.477612  5.970149 11.940299 19.402985
##       340       360       380       400       420       440       460
## 11.940299 13.432836 10.447761  5.970149  2.985075  1.492537  1.492537
```

```
### 2016
(psd.16.freq20 <- xtabs(~lcat20,data=psd.16))
```

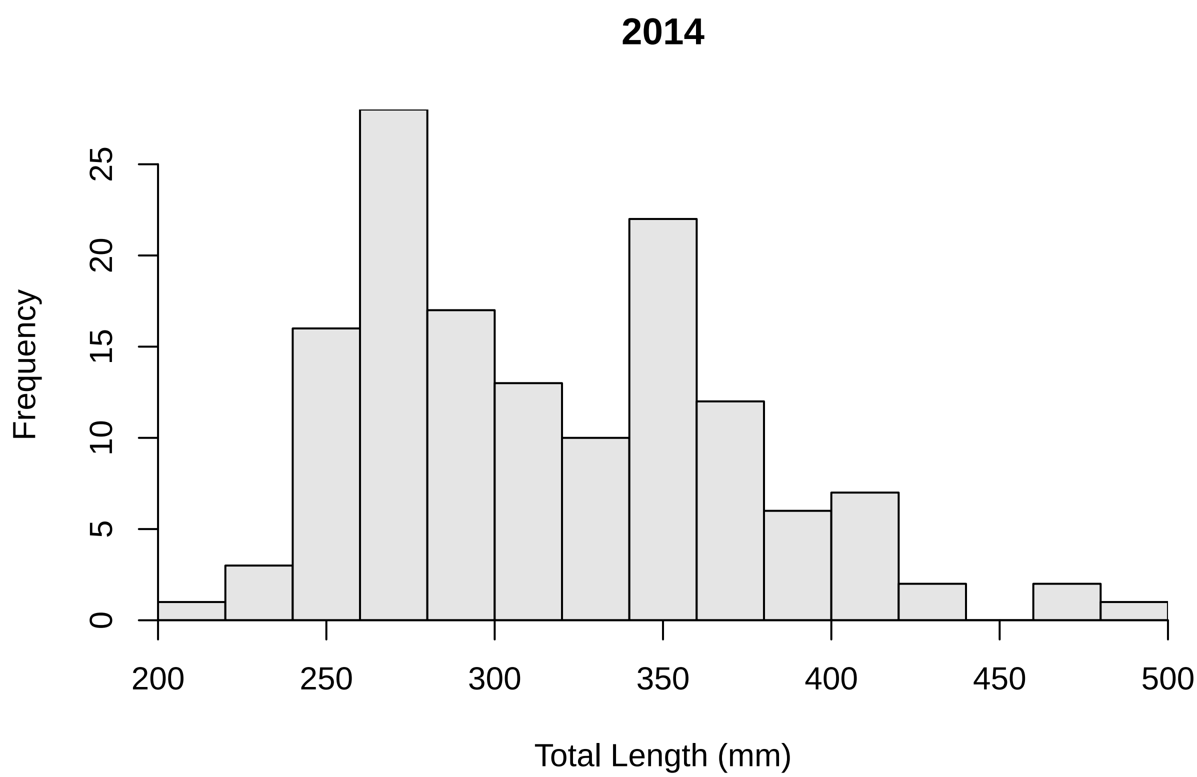
```
## lcat20
## 200 220 240 260 280 300 320 340 360 380 400 420
##   5   8  10  15  14  12   7  12  13   7   2   2
```

```
### calculate the percentage of the fish in each length nin
prop.table(psd.16.freq20)*100
```

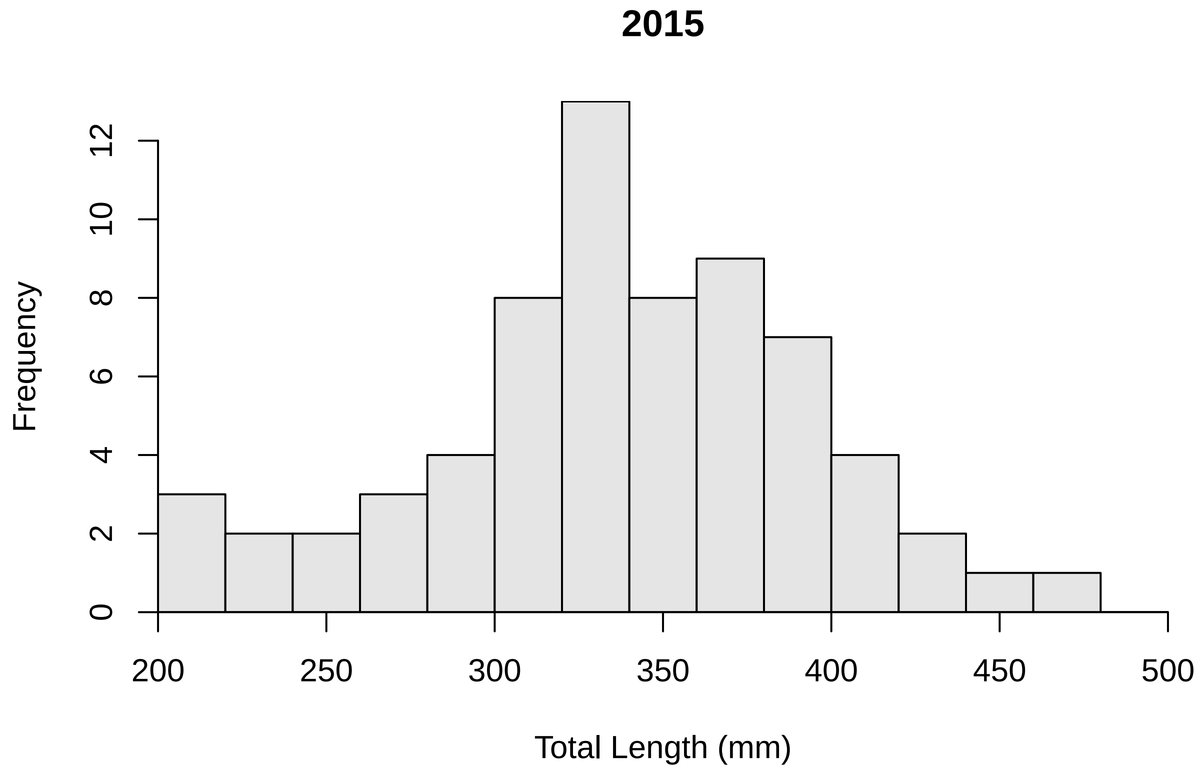
```
## lcat20
##       200       220       240       260       280       300       320
##  4.672897  7.476636  9.345794 14.018692 13.084112 11.214953  6.542056
##       340       360       380       400       420
## 11.214953 12.149533  6.542056  1.869159  1.869159
```

Lets view a quick histogram of the frequency of fish in each length bin.

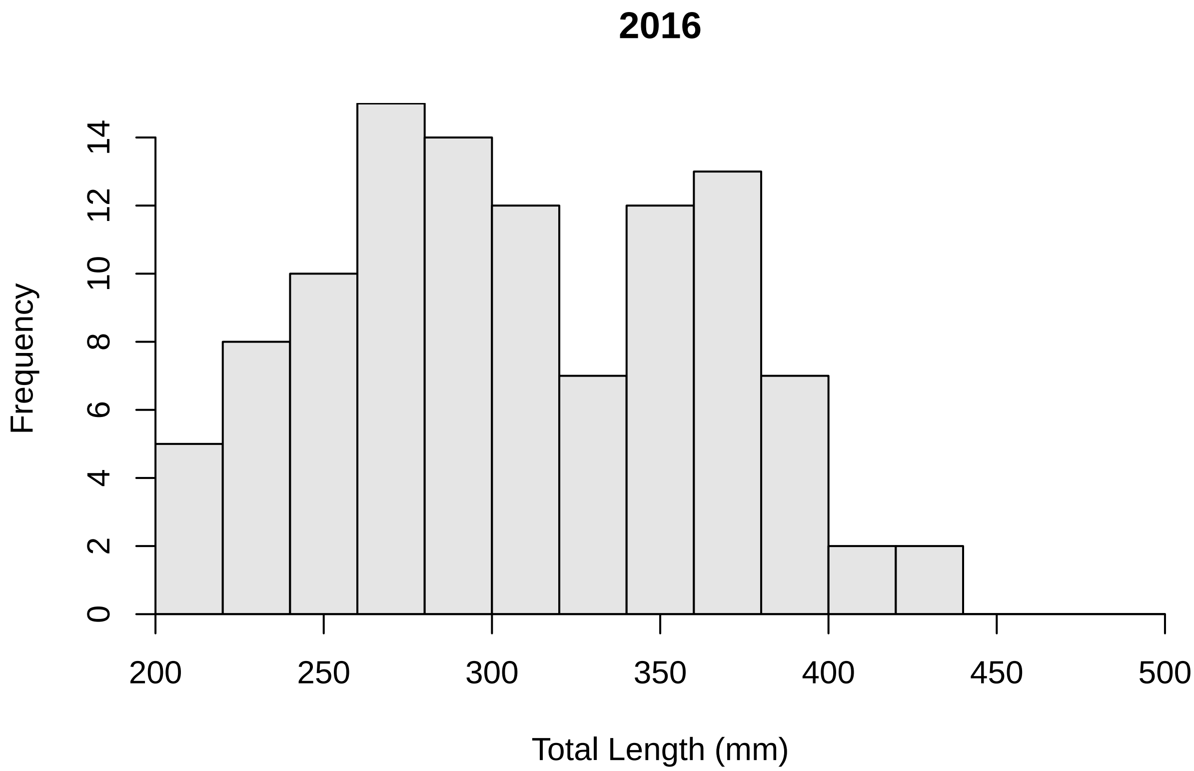
```
par(mfrow=c(3,1)) ### arrange graphs 1 column 3 rows
par(mfrow=c(1,1)) ### reset to default

### 2014
hist(~Length,data=psd.14,
     breaks = seq(200,500,20),
     main="2014",
     xlab="Total Length (mm)")
```



```
### 2015
hist(~Length,data=psd.15,
     breaks = seq(200,500,20),
     main="2015",
     xlab="Total Length (mm)")
```




```
### 2016
hist(~Length,data=psd.16,
     breaks = seq(200,500,20),
     main="2016",
     xlab="Total Length (mm)")
```

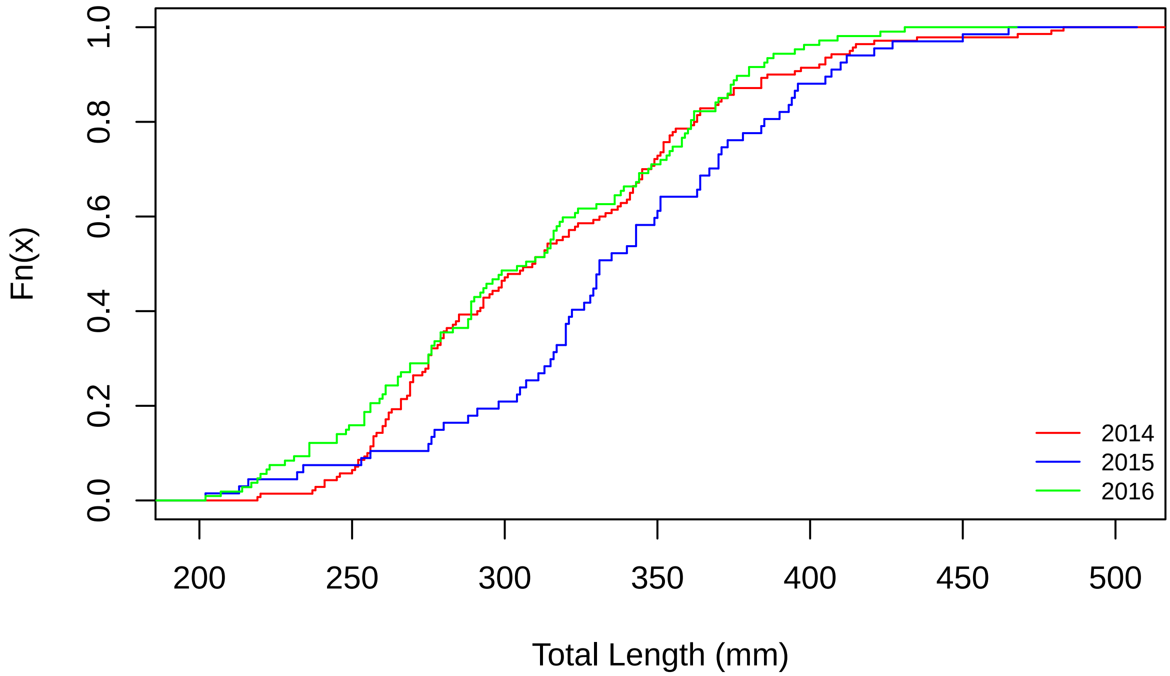


Based on the length frequency histograms above I would say that the largemouth bass population in lake Erie appears to be stable. The above graphs depict a even and gradual devrease in the number of individuals with length.

Cumulative Frequencies

Lets look at the empiricle cumulative distribution function (ECDF). This is the proportion of fish less than each observed length. This should help me compare the length frequency distributions between years.

```
par(mfrow=c(1,1)) ### reset to default

### 2014
plot(ecdf(psd.14$Length),xlab="Total Length (mm)", do.points=FALSE,
     verticals=TRUE, main="", col.01line = NULL, col="red")
### 2015
plot(ecdf(psd.15$Length),xlab="Total Length (mm)", do.points=FALSE,
     verticals=TRUE, main="", col.01line = NULL, col="blue", add=T)
### 2016
plot(ecdf(psd.16$Length),xlab="Total Length (mm)", do.points=FALSE,
     verticals=TRUE, main="", col.01line = NULL, col="green", add=T)

legend("bottomright", c("2014","2015","2016"), col=c("red","blue","green"),
       lty=1, bty="n", cex=0.75)
```



It looks to me like there are about the same length distributions each year. 2015 seems to have fewer smaller fish, however 2016 and 2014 look almost identical. This apears to support my prior assertion that the largemouth bass population in Lake Erie is stable.

Proportional Size Distribution (PSD)

Notes from Guy and Brown 9.5.2 Stock Density Indices

I will look at some size distribution indecies in order to supplement the above length frequency analysis. I can use these to test for a correlation between size structure and other factors (However I only have 3 years of data). A question I should ask is whether the index value (size structure) reflects the density and dynamics of a fish population (Willis et al. 1993). As density increases PSD tends to decrease; declines in size structure can be attributed to slowing of growth and increased mortality as resources become scarce. However, a low PSD may also occur at low population densities due to overharvest or poor habitat (Guy and Brown CH9 pg 413).

As Growth Increases there is a tendancy for PSD to increase. Low densities may result in high growth whereas high densities may result in slow growth (Guy and Brown CH9 pg 413).

Several studies have demonstrated that body condition is positively correlated to growth rate (See chapter 10 Guy and Brown). Individuals from low-density populations in which PSD in high tend to have high body condition values, and individuals from high-density populations in which PSD is low tend to have low body condition values. However body condition is an intantaneous measure, and slow growing fish may exhibit high body conditions at times of the year when food is abundant or when gonads are mature during the spawning period (Guy and Brown CH9 pg 413).

When total anual mortality increases there is a tendancy for PSD to decrease.High mortality due to overharvest and poor habitat also result in low PSD (Guy and Brown CH9 pg 414).

Predator and prey PSD are inversley related, however, the likliehood of this inverse relationship declines with the size of the water body. Carline et al. (1984) suggests that in Ohio impoundments, inverse relationships between size structure of largemouth bass and bluegill may not be expected in impoundments greater than 15 ha in size (Guy and Brown CH9 pg 414).

Note I may want to repete these analysis using CPUE instead of individuals

Lets get started with PSD

Lets get started calculating Proportional Size Distributions (PSD). I mentioned earlier using the Prefered and Quality lengths, however, I think I will also look at the relative PSD of 457 mm fish prefered by anglers. This may not work though since I think these individuals are rare in my sample and perhaps absent in some years.

Lets start by looking at the frequency of fish in each gabelhouse length category.

```
### 2014
(gfreq.14 <- xtabs(~gcat,data=psd.14))
```

```
## gcat
## preferred   quality     stock
##        18        57        65
```

```
### Convert freq to percentage
(psdXY1.14 <- prop.table(gfreq.14)*100)
```

```
## gcat
## preferred   quality     stock
##  12.85714  40.71429  46.42857
```

```
### 2015
(gfreq.15 <- xtabs(~gcat,data=psd.15))
```

```
## gcat
## preferred   quality     stock
##        15        38        14
```

```
### Convert freq to percentage
(psdXY1.15 <- prop.table(gfreq.15)*100)
```

```
## gcat
## preferred   quality     stock
##  22.38806  56.71642  20.89552
```

```
### 2016
(gfreq.16 <- xtabs(~gcat,data=psd.16))
```

```
## gcat
## preferred   quality     stock
##        11        44        52
```

```
### Convert freq to percentage
(psdXY1.16 <- prop.table(gfreq.16)*100)
```

```
## gcat
## preferred   quality     stock
##  10.28037  41.12150  48.59813
```

Note I may want to compare incrimental PSD indices between years to look for changes in fish between Quality and Prefered length or prefered and Relative length of 457mm. I should research what the advantages of incrimental PSD inices are in greater detail first.

```
### first calculate the percentage of fish length x and above
### 2014
(psdx.14 <- cumsum(psdXY1.14))
```

```
## preferred   quality     stock
##  12.85714  53.57143 100.00000
```

```r
### 2015
(psdx.15 <- cumsum(psdXY1.15))
```

```
## preferred   quality     stock
##  22.38806  79.10448 100.00000
```

```r
### 2016
(psdx.16 <- cumsum(psdXY1.16))
```

```
## preferred   quality     stock
##  10.28037  51.40187 100.00000
```

I'm having a problem here and I'm not sure what is going on. The order od my variables (preferred, quality, stock) is given in alphebetical order and is opposit that in the book. using rcumsum gives prefered PDS of 100. So I use cume sum which apears to work right. However, this is a problem when estimating the CI below. Maybe I just need to reverse the 0s and 1s in the psdCI() function. Meaning psdCI(c(p,q,s)) instead of psdCI(c(s,q,p)).

```r
### find the # of levels in gcat
levels(psd.14$gcat)
```

```
## [1] "preferred" "quality"   "stock"
```

```r
levels(psd.15$gcat)
```

```
## [1] "preferred" "quality"   "stock"
```

```r
levels(psd.16$gcat)
```

```
## [1] "preferred" "quality"   "stock"
```

```r
### 3 levels in gcat... so CI for PSD-Q = c(1,1,0), PSD-P = c(1,0,0) I think

### Make matrix of values to quickly compute CI for PSD-Q and PSD-P
(ivmat <- rbind("PSD-Q"=c(1,1,0),
                "PSD-P"=c(1,0,0)))
```

```
##       [,1] [,2] [,3]
## PSD-Q    1    1    0
## PSD-P    1    0    0
```

```r
### 2014 ###########################################
### Compute CI for multiple Indices
psdXY2.14 <- t(apply(ivmat,FUN = psdCI,MARGIN = 1,
                     ptbl=psdXY1.14,n=sum(gfreq.14),
                     method="multinomial"))
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.

## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
## Warning: Category sample size (18) <20, CI coverage may be lower than 95%.
```

```r
colnames(psdXY2.14) <- c("Estimate", "95% LCI", "95% UCI")
psdXY2.14
```

```
##       Estimate 95% LCI 95% UCI
## PSD-Q     53.6    43.3    63.9
```

```
## PSD-P     12.9     5.9    19.8
```

```
### Individual PSD indices
psdCI(c(1,1,0),ptbl=psdXY1.14, n=sum(gfreq.14), method = "binomial",
      label = "PSD-Q")
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
##        Estimate 95% LCI 95% UCI
## PSD-Q     53.6    45.3    61.6
```

```
### Not sure what is going on here the estimate form the CI is different than the estiamte in psdx.14
### Reversing 0s and 1s fixed this... psdCI(c(p,q,s)) instead of psdCI(c(s,q,p))
### Also getting the warning below
### Warning message:
### 'ptbl' not a table of proportions; attempted to convert
###  to proportions to continue.
psdCI(c(1,0,0),ptbl=psdXY1.14, n=sum(gfreq.14), method = "binomial",
      label = "PSD-P")
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
##        Estimate 95% LCI 95% UCI
## PSD-P     12.9     8.3    19.4
```

```
### 2015 ##############################################
### Compute CI for multiple Indices
psdXY2.15 <- t(apply(ivmat,FUN = psdCI,MARGIN = 1,
                     ptbl=psdXY1.15,n=sum(gfreq.15),
                     method="multinomial"))
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
## Warning: Category sample size (15) <20, CI coverage may be lower than 95%.
```

```
colnames(psdXY2.15) <- c("Estimate", "95% LCI", "95% UCI")
psdXY2.15
```

```
##        Estimate 95% LCI 95% UCI
## PSD-Q     79.1    66.9    91.3
## PSD-P     22.4     9.9    34.9
```

```
### Individual PSD indices
psdCI(c(1,1,0),ptbl=psdXY1.15, n=sum(gfreq.15), method = "binomial",
      label = "PSD-Q")
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
##        Estimate 95% LCI 95% UCI
## PSD-Q     79.1    67.9    87.1
```

```
psdCI(c(1,0,0),ptbl=psdXY1.15, n=sum(gfreq.15), method = "binomial",
      label = "PSD-P")
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
```

```
##  to proportions to continue.
```

```
##       Estimate 95% LCI 95% UCI
## PSD-P     22.4    14.1    33.7
```

```
### 2016 ###################################
### Compute CI for multiple Indices
psdXY2.16 <- t(apply(ivmat,FUN = psdCI,MARGIN = 1,
                     ptbl=psdXY1.16,n=sum(gfreq.16),
                     method="multinomial"))
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
## Warning: Category sample size (11) <20, CI coverage may be lower than 95%.
```

```
colnames(psdXY2.16) <- c("Estimate", "95% LCI", "95% UCI")
psdXY2.16
```

```
##       Estimate 95% LCI 95% UCI
## PSD-Q     51.4    39.6    63.2
## PSD-P     10.3     3.1    17.5
```

```
### Individual PSD indices
psdCI(c(1,1,0),ptbl=psdXY1.16, n=sum(gfreq.16), method = "binomial",
      label = "PSD-Q")
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
##       Estimate 95% LCI 95% UCI
## PSD-Q     51.4      42    60.7
```

```
psdCI(c(1,0,0),ptbl=psdXY1.16, n=sum(gfreq.16), method = "binomial",
      label = "PSD-P")
```

```
## Warning: 'ptbl' not a table of proportions; attempted to convert
##  to proportions to continue.
```

```
##       Estimate 95% LCI 95% UCI
## PSD-P     10.3     5.8    17.5
```

Lets try the other way of calculating CI for PSD.

```
### 2014
 ## Traditional
psdCalc(~Length,data=psd.14,species = "Largemouth Bass",what="traditional")
```

```
## Warning: Some category sample size <20, some CI coverage may be
##  lower than 95%.
```

```
##       Estimate 95% LCI 95% UCI
## PSD-Q       54      43      64
## PSD-P       13       6      20
```

```
 ## Incremental
psdCalc(~Length,data=psd.14,species = "Largemouth Bass",what="incremental")
```

```
## Warning: Some category sample size <20, some CI coverage may be
```

```
##  lower than 95%.

##          Estimate 95% LCI 95% UCI
## PSD S-Q       46      36      57
## PSD Q-P       41      31      51
## PSD P-M       13       6      20
```

```
### 2015
 ## Traditional
psdCalc(~Length,data=psd.15,species = "Largemouth Bass",what="traditional")
```

```
## Warning: Some category sample size <20, some CI coverage may be
##  lower than 95%.

##        Estimate 95% LCI 95% UCI
## PSD-Q       79      67      91
## PSD-P       22      10      35
```

```
 ## Incremental
psdCalc(~Length,data=psd.15,species = "Largemouth Bass",what="incremental")
```

```
## Warning: Some category sample size <20, some CI coverage may be
##  lower than 95%.

##          Estimate 95% LCI 95% UCI
## PSD S-Q       21       9      33
## PSD Q-P       57      42      72
## PSD P-M       22      10      35
```

```
### 2016
 ## Traditional
psdCalc(~Length,data=psd.16,species = "Largemouth Bass",what="traditional")
```

```
## Warning: Some category sample size <20, some CI coverage may be
##  lower than 95%.

##        Estimate 95% LCI 95% UCI
## PSD-Q       51      40      63
## PSD-P       10       3      17
```

```
 ## Incremental
psdCalc(~Length,data=psd.16,species = "Largemouth Bass",what="incremental")
```

```
## Warning: Some category sample size <20, some CI coverage may be
##  lower than 95%.

##          Estimate 95% LCI 95% UCI
## PSD S-Q       49      37      60
## PSD Q-P       41      29      53
## PSD P-M       10       3      17
```